

UC Berkeley

UC Berkeley Previously Published Works

Title

Parameter variability and the interpretation of physiologically based pharmacokinetic modeling results.

Permalink

<https://escholarship.org/uc/item/93z4z4tk>

Journal

Environmental health perspectives, 102 Suppl 11(Suppl 11)

ISSN

0091-6765

Authors

Spear, RC
Bois, FY

Publication Date

1994-12-01

DOI

10.1289/ehp.94102s1161

Peer reviewed

Parameter Variability and the Interpretation of Physiologically Based Pharmacokinetic Modeling Results

Robert C. Spear¹ and Frederic Y. Bois²

¹School of Public Health, University of California, Berkeley, California; ²Lawrence Berkeley Laboratory, Berkeley, California

For the past several years we have been working with models of benzene distribution and metabolism, principally in the rat, but more recently in humans. Our biologically related objectives have been primarily to assist our laboratory-based colleagues in their quest for understanding of the mechanisms by which benzene exerts its toxic action. A secondary goal has been to develop or adapt models useful in risk assessment applications. We have also had methodological goals that relate to applications of sensitivity analysis on the one hand, but more fundamentally to the connection between experimental data and model structure and parameterization. This paper presents an overview of our work in these areas — Environ Health Perspect 102(Suppl 11):61–66 (1994)

Key words: modeling, benzene distribution, benzene metabolism, rats, human

Introduction

The postulate underlying much of our work is that the construction and interpretation of mathematical models of biological systems must recognize the population variability inherent in the processes under analysis. In our work, this has been realized in two ways. We assign statistical distribution functions to the constant parameters of the model to reflect population variability in their values. Second, in much of our work we judge the adequacy of a particular simulation not by a continuous measure of goodness of fit to laboratory or field data, but by a binary outcome index which reflects our judgment that the simulation is or is not adequately representative of the data. There is obviously a subjective aspect to the judgment of the adequacy of a particular simulation, but we feel that the benefits gained outweigh the costs.

Because many of the generally interesting results of our work arise from this basic analytical procedure, it is necessary to give a somewhat more complete explanation of the process through example, as well as to give some further practical justification of its features. The first model that we

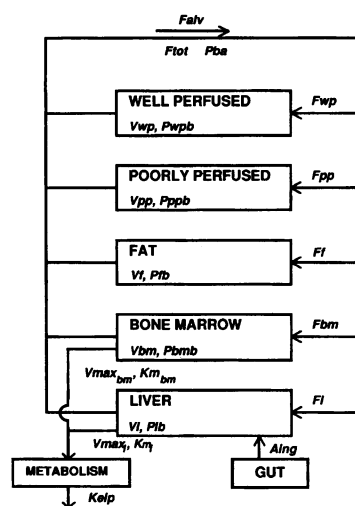


Figure 1. Schematic representation of the physiologic model used to simulate the distribution of benzene. Symbols: V, volumes; F, flows; P, partition coefficients; V_{max} , K_m , Michaelis–Menten parameters; K_{elp} , metabolite excretion rate; A_{ing} , absorption rate.

worked with extensively described the distribution of benzene in Fischer rats. It is a five-compartment physiologically based pharmacokinetic (PBPK) model with Michaelis–Menten terms describing the transformation of benzene into its first metabolites in the liver and the bone marrow. The general structure is shown in Figure 1.

In our analysis of this model, we chose ranges that define the limits of uniform distributions for each of the parameters or their scaling variables. To take into account known structural dependencies

between the parameters, we made many of them a function of body weight. This was accomplished by defining three body weight scaling functions, each of which contained a scaling coefficient that was assigned a range; this was also the case for the unscaled model parameters. The ranges for all 26 variables were selected to include values reported in the literature for adult rats. Monte Carlo simulations were used to study the effects of variability in the model parameters on its output. The parameters were sampled uniformly or log-uniformly from their ranges.

The model was evaluated against data from three different sets of experiments in what one might term a mini-metaanalysis (1). All three experiments were performed on male Fischer-344 rats. In the first set of experiments, carried out by Rickert et al. (2), rats were exposed to benzene through inhalation. These experiments provide extensive information about the kinetics of benzene in blood, fat, bone marrow, and exhaled air. The tissue concentrations were recorded both during and after exposure. Three rats were placed in an inhalation chamber for a series of exposure intervals. At the end of each exposure interval, the rats were killed, and the concentration of benzene in the fat, blood, and bone marrow was measured.

The second two sets of experiments were performed by Sabourin et al. (3,4). In one set, rats were exposed to benzene through inhalation. For each dosing level, they were exposed by nose only for 6 hr, and the total amount of benzene exhaled

This article was presented at the Workshop on Pharmacokinetics: Defining the Dose for Risk Assessment held 4–5 March 1992 at the National Academy of Sciences in Washington, DC.

This work was supported by National Institutes of Health grant P42 ES04705.

Address correspondence to Dr. Robert C. Spear, School of Public Health, University of California, Berkeley, CA 94720. Telephone (510) 642-0761. Fax (510) 642-5815.

and the total amount of metabolites excreted were obtained for the following 56 hr. In the other set of experiments, the rats were exposed through gavage also at various dosing levels. For each dosing level, the total amount of benzene exhaled and the total amount of metabolites excreted were measured for the following 48 hr.

Before turning to the mini-metanalysis, it is instructive to consider first the kind of information given by the Monte Carlo experiments in contrast to the more usual analysis in which most of the model parameters are fixed at biologically plausible values before a few important parameters are adjusted to give a best fit to the experimental data, typically those dealing with metabolism. We carried out such a study (5) using the results of the benzene modeling work of Travis et al. and Medinsky et al. (6–8) to compare to our own results, in which we employed parameter distributions rather than fixed values. Our model was structurally similar to the Travis and Medinsky models (their parameter values were always within the ranges used in our Monte Carlo simulations). The comparisons used the log-likelihood function as a continuous measure of goodness of fit between the model predictions and the experimental data points for the Travis model and parameter set; for that of Medinsky; and then for each of 1000 Monte Carlo simulations for our own model. Figure 2 shows a plot in which each experimental data point is plotted against its simulated value for all the data in the Rickert experiments. Not surprisingly, the best Monte Carlo run gives a better overall fit than either the Travis or Medinsky models, since the best Monte Carlo fit is, in effect, the result of an optimizing search. The same outcome can be seen for the Sabourin inhalation and gavage data in Figure 3.

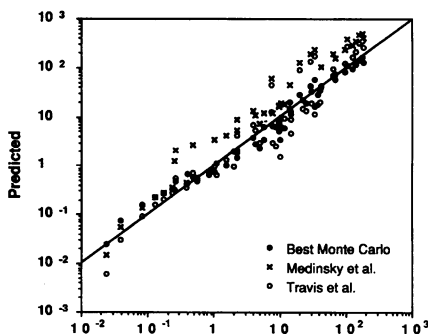


Figure 2. Predictions made using three published models, versus observed values of the data points obtained during Rickert et al. (2) experiments. For a perfect fit, the points would fall on the identity line. The best Monte Carlo simulation lies the closest to the identity line.

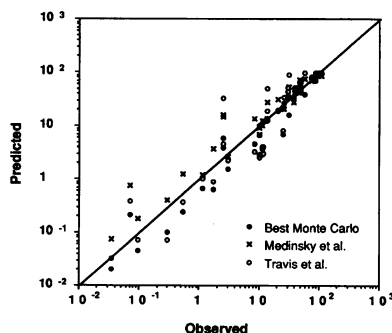


Figure 3. Predictions made using three published models, versus observed values of the data points obtained during Sabourin et al. (3,4) experiments. For a perfect fit, the points would fall on the identity line.

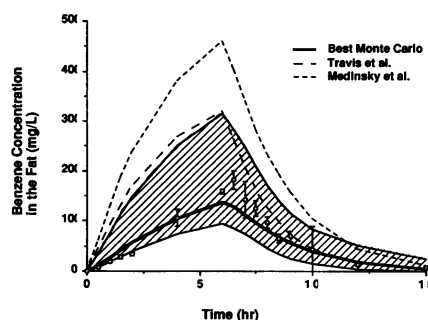


Figure 4. Benzene concentration in the fat of Fischer-344 rats, during and after a 6 hr exposure to 490 ppm benzene in the air. The open circles represent the experimental data (\pm SD) of Rickert et al. (2). The time profiles obtained using three published models are given. The shaded area encloses the results of the 200 Monte Carlo simulations of highest log-likelihood.

Somewhat more instructive are the next figures showing the time course of several measured variables. Figure 4 shows benzene in the fat for the Rickert data, where the shaded area encloses the trajectories of the 200 best Monte Carlo runs, all of which had lower log-likelihoods than either the Travis or Medinsky models. Similarly, Figure 5 shows the benzene concentration in the venous blood for the Rickert experiment. It can be seen that the Travis model shows a much better fit than the best Monte Carlo simulation, which points out that the best fit overall, as measured by the log-likelihood function, does not guarantee a best or even a good fit to a particular subset of the total data set. Finally, Figure 6 shows the percentage of the administered dose excreted as metabolic products during the 48 hr following benzene gavage in the Sabourin experiment. Here, all three models give similar results, presumably because Travis and Medinsky paid particular attention to fitting these data in their adjust-

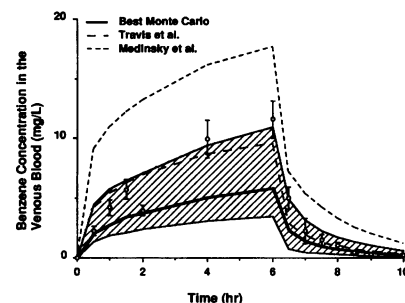


Figure 5. Benzene concentration in the venous blood of Fischer-344 rats, during and after a 6 hr exposure to 490 ppm benzene in the air. The open circles represent the experimental data (\pm SD) of Rickert et al. (2). The time profiles obtained using three published models are given. The shaded area encloses the results of the 200 Monte Carlo simulations of the highest log-likelihood.

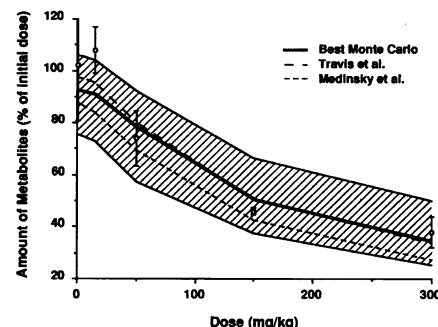


Figure 6. Percentage of the administered dose excreted as metabolites by Fischer-344 rats, during the 48 hr following benzene gavage. The open circles represent the experimental data (\pm SD) of Sabourin et al. (3,4). The profiles obtained using three published models are given. The shaded area encloses the results of the 200 Monte Carlo simulations of highest log-likelihood.

ments of the metabolic parameters K_m and V_{max} .

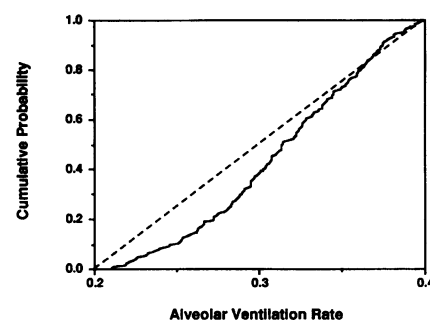
The Pass-Fail Criterion of Fit

A worrying feature inherent in the foregoing studies is the implicit reliance placed on the mean values of very small numbers of observations per data point. Given the variability to be expected in these mean values, it is important to keep in mind that the measure of goodness of fit of the model to these data is itself a random variable of unknown distribution. In past studies, this same uncertainty has led us to choose an alternative to a continuous measure of goodness of fit. Our alternative is to formulate a set of criteria which attempt to define the adequacy of a simulation as either having captured the essence of the data or not having done so. We have, in effect, explored the notion of a binary criterion of

Table 1. Definition of additional criteria used for evaluating the results for Rickert's experiments.

Variable	Observation time	Range, %	Reference time, hr
(During 8-hr exposure)			
Bone marrow concentration	0.5	15.0–55.0	6.0
Fat concentration	2.0	5.0–35.0	6.0
Venous blood concentration	1.5	30.0–70.0	6.0
(After 6-hr exposure)			
Bone marrow concentration	6.5	20.0–60.0	6.0
	9.0	0.02–5.0	6.0
Fat concentration	8.0	30.0–70.0	6.0
	15.0	1.0–10.0	6.0
Venous blood concentration	6.5	30.0–70.0	6.0
	10.0	1.0–10.0	6.0
Quantity remaining to be exhaled	7.0	25.0–65.0	6.5
	15.0	1.0–5.0	6.5

Values were computed for each of the appropriate compartments following a 6- or 8-hr exposure to 490 ppm of benzene. To be a *pass*, the value of each variable at the given observation times must be within the given range (%) of its value at the reference time (e.g., the stimulated bone marrow concentration at 0.5 hr must be within 20 to 60% of its value at 6 hr). Time is counted from the start of exposure.

**Figure 7.** Cumulative probability plots for pass (solid) and fail (dash) for the alveolar ventilation rate parameter from Sabourin inhalation experiment.

Wilcoxon test rejects the hypothesis that the passing and failing distributions are identical at the 5% level. The conclusion is that, within the range defined by the prior distribution, alveolar ventilation rate is a parameter important in determining successful from unsuccessful simulations. Over half of the parameters were important by this criterion for the Rickert experiments versus less than one-third for the two Sabourin experiments. In retrospect, this difference might have been expected given the greater number of variables measured by Rickert.

An interesting outcome of this analysis was that, despite each of the parameter distributions extending over what were thought to be biologically plausible ranges, we found that passing simulations were not easy to achieve. After some minor distribution trimming at the outset, we were only able to reach 6.6% passes in the Rickert experiment, 11% in Sabourin inhalation, and 18.6% in Sabourin gavage. However, passing simulations were observed over the entire range of parameter values allowed by the prior distributions. This was true of sensitive, as well as insensitive, parameters as classified by the univariate criterion. This led us to the conclusion that the parameter space is highly structured, with considerable interaction between parameters. Clearly, multivariate analysis was in order.

Multivariate Analysis

Only one of the various multivariate analyses that we carried out will be described because it goes directly to the metanalysis idea. In particular, if the model structure and parameter distributions were completely consistent over all three data sets, then a statistical analysis of passing parameter vectors should be able to discern no differences in those multivariate distributions by experiment. That is, a parameter vector

goodness of fit as a way to avoid overreliance on highly variable experimental data, and we now turn to the application of that idea in the context of our analysis of the Rickert and Sabourin experiments.

Because different variables were measured in each of these experiments at different times, it was clear that the criteria for a successful simulation run would be specific to each experiment. In general, our criteria were stated as acceptable ranges about the published values of the measured variables at the particular times the investigator chose to make the measurement. Table 1 shows a subset of these criteria for the Rickert experiment. Sixteen criteria were used for Rickert, 15 for Sabourin inhalation, and 16 for Sabourin gavage. In each case, if one of the criteria was met, a 1 was assigned (or a 0 if criteria were not met). Therefore, for each of the three experimental data sets, N Monte Carlo simulations produced a (0,1) pass-fail matrix of P (either 15 or 16) columns and N rows. To be classified as pass overall, a simulation had to pass all the criteria for a given experiment (i.e., for the k th simulation to be a pass, all P entries in the k th row must equal 1). Clearly, this procedure avoids the problem noted above when minimizing the log-likelihood measure of goodness of fit in that each measurement subset of the data must lie within acceptable bounds; a very good fit to one subset cannot compensate for a bad fit to another.

Posterior Analysis

Given the model structure, the parameter distributions, and the outcome classification algorithm, one can now carry out Monte Carlo runs for either the Rickert or the

Sabourin experiments by randomly selecting a parameter vector, carrying out the simulation, and classifying the result as a pass or a fail by reference to the predetermined classification rules. Repetition of this process n times leads to m passing parameter vectors and $n - m$ failing vectors. The central feature of the procedure is the posterior analysis of the parameter vectors using statistical procedures of various sorts to determine which elements of the parameter vector appear to be responsible for passing or failing simulations. For the moment, assume that a sufficient number of passes are obtained to make this a feasible proposition.

The first step in the posterior analysis is to inspect the univariate marginal distributions. If $F(x_i)$ is the cumulative distribution of the prior, then we ask if the cumulative distributions of this parameter differ between passing and failing simulations (i.e., does $F(x_i | p) = F(x_i | f)$ where p stands for pass and f for fail?). The contention is that if differences can be detected by some suitable hypothesis test, then x_i is a sensitive parameter. If no difference can be detected, then multivariate analysis is needed to further elucidate potential interactions in the parameter space. (Clearly, statistical power is also an issue, but in Monte Carlo work, this is usually an issue of cost arising from model run time considerations.) Figure 7 shows the cumulative distribution of alveolar ventilation rate for the Sabourin inhalation data for the prior (the dotted line) and for passing simulations (the solid line). In these experiments the number of passing simulations was low, so the prior and the failing distributions are very similar. In this particular case, the

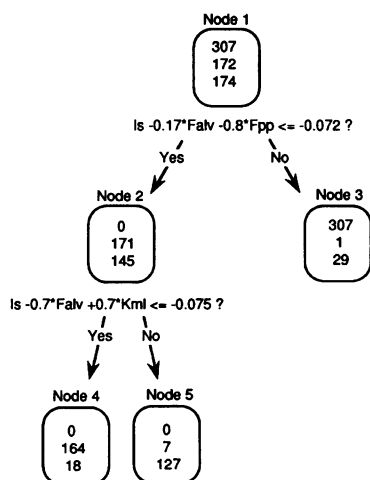


Figure 8. CART diagram from three experiments. Node 1 contains just passes, 307 from the Rickert experiment, 172 from the Sabourin inhalation experiment, and 174 from the Sabourin gavage experiment. After splitting on parameters F_{alv} and F_{pp} , Node 2 contains no Rickert passes. A split using F_{alv} and K_{ml} leads to Node 4 containing most of the Sabourin inhalation passes and Node 5 containing most of the Sabourin gavage passes.

giving rise to a successful simulation of the Rickert data should give rise to a successful simulation of the Sabourin experiments. We subjected the passing parameter vectors from the three experiments to analysis by a procedure called Classification and Regression Trees (CART). CART a computer-intensive multivariate analysis program which, in this application, attempts to find rules based on the parameter values which associate the parameter vector with a particular experiment.

Figure 8 shows the outcome of the CART run in which the program was challenged with finding parameter-based rules for classifying a vector as coming from the Rickert, Sabourin gavage, or Sabourin inhalation experiments. The analysis begins at node 1 with 307 Rickert passing vectors, 172 Sabourin inhalation passes and 174 Sabourin gavage passes. The first rule the program determined was based on a linear combination of alveolar ventilation rate and blood flow through the poorly perfused tissue, F_{pp} . (It is interesting to note that F_{pp} was not identified as an important variable in any of the univariate analyses.) If the criterion was not met, then all 307 Rickert vectors were sent to node 3 with 1 Sabourin inhalation and 29 Sabourin gavage vectors. Clearly, these two parameters alone identified the Rickert vectors with very low misclassifications of the Sabourin vectors. Most of the Sabourin vectors were sent to

node 2, where a second rule was determined based again on alveolar ventilation, but this time, coupled with the K_{ml} value in the liver. This rule does a good job of separating the Sabourin inhalation vectors, node 4, from the gavage vectors, node 5.

The final outcome of this multivariate analysis was that, on the basis of only three of the 26 parameters, CART could determine portions of the parameter space preferentially associated with each of the experiments. Various explanations spring to mind, but we note that one inhalation experiment used exposure chambers and the other a nose-mask administration system. Also, the Rickert and Sabourin experiments were performed at different altitudes and other aspects of the protocols and methods differed. Nevertheless, the ease with which CART was able to discern that differences existed was disconcerting. Clearly, the model is not consistent across these three sets of experimental data.

Methodological versus Biological Accomplishments

At this point it is interesting to reflect on what had been accomplished in this project in late 1990, when much of the foregoing work was completed. In particular, it is interesting to contrast the methodological viewpoint with the biological. In the former area, we were all surprised at how few passing simulations we obtained, particularly in view of the fact that we were dealing with a PBPK model with the advantage of assigning to the parameter distributions only what the literature suggested to be biologically plausible values. Second, it seemed curious that passing simulations extended to virtually all corners of the parameter hypercube. There was no suggestion of what we called "raisins in the pudding," isolated areas in the parameter space with a high density of passes. Rather, the results suggested a n -dimensional ribbon winding around in the parameter space. We have seen a very similar picture in a large ecosystem model that we are working with, as well as in a model of mosquito population dynamics in the context of arboviral disease transmission. It is somewhat difficult to discern anything definitive about the shape of the passing region in 26 dimensions with the very sparse number of points available, but it is clear that there must be a great deal of interaction between parameters if a particular parameter can take on any value in its range yet still achieve a pass if other parameters are in the right place. We are only now beginning to mount a more sophisticated approach to investigating these com-

plicated interactions in high-dimensional spaces. If one were to have a model that would extend from distribution and metabolism through to cellular damage, an understanding of these high-dimensional interactions would hold the key to understanding the importance of exposure-related parameters versus genetically determined parameters in the production of chemically related disease.

Turning to the interests of our laboratory-based colleagues, the situation was very different. They had little interest in the work described above. Indeed, much of this work had been concerned with the development of models that had modest potential to address issues of interest to them. Their interest is in mechanism of toxic action and it was clear that, at the very least, we had to be able to model metabolic processes to stimulate any interest at all. The stimulus to develop such a model arose in the context of the carcinogenic properties of benzene versus phenol. Benzene is pan-carcinogenic in animals, while phenol did not exhibit carcinogenic effects during its National Toxicology Program (NTP) study (9). However, phenol is on the pathway from benzene to some of the suspected ultimate carcinogens (hydroquinone and benzoquinone). On the basis of that argument, phenol should be a carcinogen, unless extensive first-pass effects prevent it from entering the general blood circulation. The challenge was to produce a model that was consistent with what was known on the distribution and metabolism of both benzene and phenol and then to determine which metabolic products showed major differences under the conditions of the NTP bioassays (gavage for benzene and drinking water exposure for phenol). We wanted, in particular, to assess whether a much larger quantity of hydroquinone had been formed during benzene exposure than during phenol exposure. Such a model was developed that tracked circulating levels of benzene and phenol and the production of the major metabolites (10). Figure 9 shows the metabolic pathways described by the model. This is the most complete model of benzene metabolism published to that time. It uses the benzene distribution model (Figure 1) and a second model describing the distribution of phenol in the body. Clearly this was a complicated model with 12 compartments and 64 parameters. Calibrating it to the pharmacokinetic and metabolic data at hand was a challenging task. We again used Monte Carlo simulations and reverted to a continuous measure of fit, log-likelihood, which has several disadvantages as men-

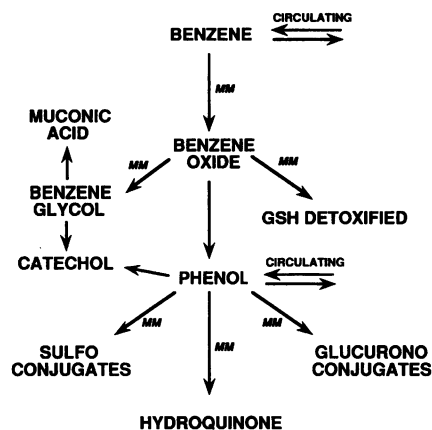


Figure 9. Our latest model of benzene metabolic pathways. MM, Michaelis-Menten reactions.

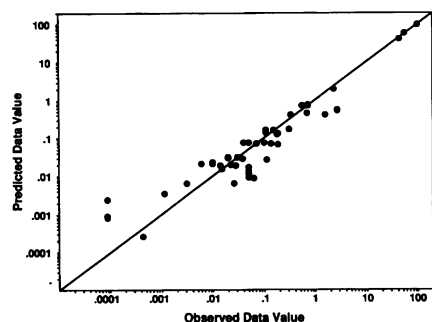


Figure 10. Predictions made by the best-fitting simulation run versus observed experimental values (3,4,11-13). For a perfect fit, the points would fall on the identity line.

tioned above, but could lead quickly to improvement in the goodness of fit. Figure 10 shows the quality of fit of the best Monte Carlo simulation (after some 3000 runs).

As shown in Figure 11, the predicted levels of phenol and hydroquinone circulating or formed during the phenol-NTP cancer bioassay were higher than during the benzene bioassay. This is in direct conflict with the hypothesis that phenol or hydroquinone are solely responsible for benzene carcinogenicity. By contrast, catechol and muconaldehyde were formed during benzene administration, but to a much lesser degree during phenol administration. We had to conclude that hydroquinone alone is unlikely to be the ultimate carcinogen. Other lines of evidence (from *in vitro* mutagenicity tests) points also to a synergy between the various metabolites.

More recently, this model was employed in a straightforward risk assessment mode to explore whether peak exposures to benzene are potentially more

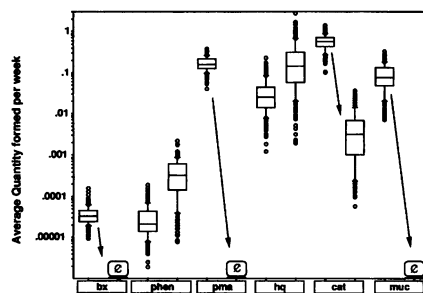


Figure 11. Predicted quantities of various metabolites formed during a week of either benzene exposure (first box of each pair, lowest gavage dose in the National Toxicology Program experiment) or phenol exposure (second boxes, highest drinking water in the NTP experiment). Symbols: bx, benzene oxide; phen, phenol (concentration in blood, not quantity formed); pma, phenyl mercapturic acid; hq, hydroquinone; cat, catechol; muc, muconaldehyde.

damaging than constant exposures in the occupational environment (11). The model was parameterized using Monte Carlo simulations as described previously. The 200 sets of parameter values giving the best fits were then used to predict the amount of metabolites formed in 24 hr after inhalation exposure to either 1 ppm for 8 hr or 32 ppm for 15 min. Note that these two exposure schedules lead to the same total inhaled dose and hence have the same 8-hr time-weighted average. Under linear behavior, as has traditionally been assumed when assessing the risk of benzene carcinogenesis, the total amount of each metabolite formed would be the same for both exposures. However, results showed that for some metabolites, including the most toxic, there is a 20% difference, on average, between the amount of metabolite produced under peak versus constant exposure conditions (Figure 12). Therefore, at low levels, benzene peak exposures potentially are more detrimental than constant ambient concentrations. At high levels, as in controlled experiments, this effect is not observed. Thus, our conclusions, most relevant to human exposure levels, were only accessible through model extrapolation and not through experimentation alone. This modeling approach can be applied to obtain correct estimates of effective exposure for various cohorts of occupationally exposed individuals.

Conclusions

Our work has had three major components: methodological issues surrounding the parameterization and use of complicated models, the use of the benzene models in the formulation and analysis of hypotheses

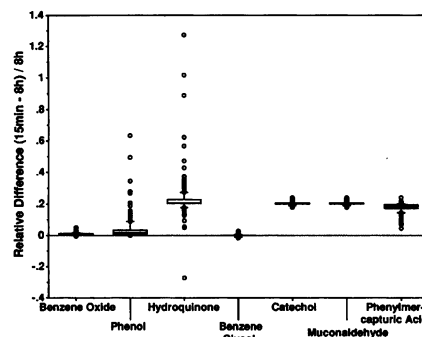


Figure 12. Box plots of the relative differences in the formation of various benzene metabolites in rats after exposures to 1 ppm for 8 hr or 32 ppm for 15 min. These results were obtained by simulation.

regarding toxic mechanisms, and the use of the benzene models in risk assessment applications. While our applications of the models have been similar to many other applications of PBPK models, our methodological approach has some unique features. The obvious question concerns what insights we may have gained that can be attributed to the use of the more complicated Monte Carlo methods we have employed. Although the point was not emphasized above, we have no doubt that the use of parameter distributions and the pass-fail fit criterion lead to a robust form of sensitivity analysis that has significant advantages in these applications over the fixed-parameter perturbation methods conventionally employed. The principal advantage of our approach is that it acknowledges that, in most biological applications, there is no clearly defensible way to select a "best" set of parameters about which to carry out traditional sensitivity analysis.

Indeed, the advantage offered by the Monte Carlo approach does seem to relate generally to the notion of robustness in avoiding overreliance on particular parameter values in the absence of very compelling and specific data defining their population distribution. In all of the foregoing analyses our conclusions rested not on a single set of model parameters, but numerous sets whose values vary extensively over biologically plausible values. Our conclusions relate to the behavior of the model over regions of the parameter space rather than at a single point. Returning to Figure 11, for example, the rejection of hydroquinone as the ultimate carcinogen is based on the fact that the hydroquinone distribution under phenol administration generally exceeds that of the corresponding distribution under benzene administration. Clearly, there are pairs of individual simu-

lations, both of which give rise to good fits to the experimental data, where the production of hydroquinone was two orders of magnitude greater in the benzene experiment than in the phenol experiment and vice versa. That is, there are particular sets of parameter values that would produce simulation results supporting either the conclusion that hydroquinone is important

or that it is not. Looked at from the perspective of the totality of acceptable simulations, however, the pathway through benzene glycol to catechol and muconaldehyde looks much more likely.

In the context of risk assessment too, there is much to be said for taking a distributional view rather than focusing on point estimates of the risk of the most exposed

individual. Here the context is somewhat different, but the underlying variability in population susceptibility or in the uncertainty of our knowledge of parameter values is just as much of a problem and has just as great a potential for leading to inappropriate risk management decisions.

REFERENCES

1. Spear RC, Bois FY, Woodruff T, Auslander D, Parker J, Selvin S. Modeling benzene pharmacokinetics across three sets of animal data: parametric sensitivity and risk implications. *Risk Anal* 11:641-654 (1991).
2. Rickert DE, Baker TS, Bus JS, Barrow CS, Irons RD. Benzene disposition in the rat after exposure by inhalation. *Toxicol Appl Pharmacol* 49:417-423 (1979).
3. Sabourin PJ, Bechtold WE, Birnbaum LS, Lucier G, Henderson RF. Differences in the metabolism and disposition of inhaled [³H]benzene by F344/N rats and B6C3F1 mice. *Toxicol Appl Pharmacol* 94:128-140 (1988).
4. Sabourin PJ, Chen BT, Lucier G, Birnbaum LS, Fisher E, Henderson RF. Effect of dose on the absorption and excretion of [¹⁴C]benzene administered orally or by inhalation in rats and mice. *Toxicol Appl Pharmacol* 87:325-336 (1987).
5. Bois FY, Woodruff TJ, Spear RC. Comparison of three physiologically-based pharmacokinetic models of benzene disposition. *Toxicol Appl Pharmacol* 110:79-88 (1991).
6. Travis CC, Quillen JL, Arms A. Pharmacokinetics of benzene. *Toxicol Appl Pharmacol* 102:400-420 (1990).
7. Medinsky MA, Sabourin PJ, Lucier G, Birnbaum LS, Henderson RF. A physiological model for simulation of benzene metabolism by rats and mice. *Toxicol Appl Pharmacol* 99:193-206 (1989).
8. Medinsky MA, Sabourin PJ, Lucier G, Birnbaum LS, Henderson RF. A toxicokinetic model for simulation of benzene metabolism. *Exp Pathol* 37:150-154 (1989).
9. NTP. Toxicology and carcinogenesis studies of benzene (CAS No. 71-43-2) in F344/N rats and B6C3F1 mice (gavage studies). National Toxicology Program. Bethesda, MD: NTP, 1986.
10. Bois FY, Smith M, Spear RC. Mechanisms of benzene carcinogenesis: application of a physiological model of benzene pharmacokinetics and metabolism. *Toxicol Lett* 56:283-298 (1991).
11. Bois FY, Paxman D. An analysis of exposure rate effects for benzene using a physiologically based pharmacokinetic model. *Regul Toxicol Pharmacol* (in press).